

# Software work in mathematical biology

Maxie Dion Schmidt

April 12, 2022

---

## Abstract

Ribonucleic acid (RNA) and similar biological molecules are easily sequenced. Cutting-edge research on RNA structures and sequencing is utilized in medical applications from cancer research and therapy, mRNA-based vaccines, and may in time grow to provide the central role of diagnostics and therapeutic treatments to remedy many uncured diseases. The biological function of RNA sequences is dictated by the physical structures they fold into. Predicting, comparing and visualizing those primary and secondary structures from their base sequence data is an important and computationally complex scientific problem. Moreover, as the cellular roles for RNA molecules continue to grow, so does the importance of gaining functional insight from structural analyses. In this technical presentation we will discuss my applied software work in mathematical biology. This work was supported working with the gtDMMB group of Professor Christine Heitsch at Georgia Tech from 2018–2022. The projects we developed have resulted in cross-platform open source software for the biology community.

We focus on two primary open source software projects updated over the last few years. The first application, RNAstructViz, is a GUI-based application written in C++ for visualizing RNA secondary structures. RNA structures can be input in several standardized text formats for base sequences. The application then displays arc diagram views of base pairings across up to three folded structures concurrently. These arc diagrams are displayed on screen using vector graphics libraries and can be exported in publication quality image formats to SVG and PNG. The second application is GTFold, a fast, scalable multicore software for predicting RNA secondary structure that is one to two orders of magnitude faster than the de facto standard programs and which achieves comparable accuracy of prediction. This software was originally written in C++ as a suite of command line utilities. Over the last few years, we have worked to develop the GTFoldPython bindings and wrapper libraries to use the functionality provided by GTFold in Python3.

**RNAstructViz software documentation:** <https://github.com/gtDMMB/RNAstructViz/wiki>

**RNAstructViz application note (2021):** <https://doi.org/10.1093/bioinformatics/btab197>

**GTFoldPython software documentation:** <https://github.com/gtDMMB/GTFoldPython/wiki>

**GTFold software:** <https://github.com/gtDMMB/gtfold>

**GTFold publication (2012):** <https://doi.org/10.1186/1756-0500-5-341>